Parallel processing in language production

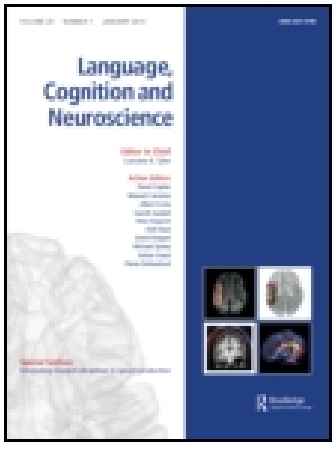
Melinger, Alissa; Branigan, Holly P.; Pickering, Martin J.

Published in:
Language, Cognition and Neuroscience

DOI:
10.1080/23273798.2014.906635

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record



Language, Cognition and Neuroscience
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/plcp21

Parallel processing in language production

Alissa Melinger[a], Holly P. Branigan[b] & Martin J. Pickering[b]
[a] School of Psychology, University of Dundee, Dundee, UK
[b] Department of Psychology, University of Edinburgh, Edinburgh, UK
Published online: 03 Apr 2014.

To cite this article: Alissa Melinger, Holly P. Branigan & Martin J. Pickering (2014) Parallel processing in language production, Language, Cognition and Neuroscience, 29:6, 663-683, DOI: 10.1080/23273798.2014.906635

To link to this article: http://dx.doi.org/10.1080/23273798.2014.906635

Language, Cognition and Neuroscience, 2014
Vol. 29, No. 6, 663–683, http://dx.doi.org/10.1080/23273798.2014.906635

Routledge
Taylor & Francis Group

Parallel processing in language production

Alissa Melinger[a]*, Holly P. Branigan[b] and Martin J. Pickering[b]

[a]*School of Psychology, University of Dundee, Dundee, UK;* [b]*Department of Psychology, University of Edinburgh, Edinburgh, UK*

(*Received 11 September 2013; final version received 17 March 2014*)

Following the Sixth International Workshop on Language Production (Edinburgh, UK, Sept., 2010), this special issue presents a collection of contributions concerned with a wide range of representational and processing components. In the present article, we review the evidence for parallel processing at different levels within the production system with the aim of identifying any generalisation or common characteristics that might underpin a robust model of language production. Our review synergises with the other articles of the special issue. After reviewing the literature, we conclude that the evidence for parallelism is stronger at some levels than at others and it is premature to take a strong stand for a unified principle that applies equally to all components of the production system. Following our review, we introduce the other articles represented within this special issue.

Keywords: language production; parallel processing; serial processing

In 2010, the University of Edinburgh and the University of Dundee jointly hosted the 6th International Workshop on Language Production. This series of workshops, founded in 2004, has as its mission statement the goal of bringing young and established production researchers together to discuss new and emerging themes, to consider the future of the field, and to present fresh perspectives on old questions. The workshop fulfilled that remit by assembling an exciting programme of talks by esteemed researchers from the old world and new (the workshop programme is available online via http://www.lang-prod.org/) and included extended discussion sessions to promote vibrant debate. The talk topics covered the gamut of the field of production, spanning behavioural and neuroscientific methods to investigate lexical, sentential and discourse processes, both from a more traditional psycholinguistic perspective and from cognitive and evolutionary viewpoints.

This special issue contributes to the aims of the workshop series by continuing the debate in these pages. We have assembled a selection of empirical and review papers by the invited speakers. Like the workshop, the papers in this special issue straddle traditional boundaries by investigating the interactions between processing stages and domains. We present papers discussing the locus of syllable frequency effects (Perret, Schneider, Dayer, & Laganaro, this issue), a comparison of the lexical selection mechanism for bound vs. free grammatical morphemes (Jescheniak, Schriefers, & Lemhoefer, this issue; Janssen, Schiller, & Alario, this issue; Jescheniak, Schriefers, & Lemhoefer, this issue_b), the considerations that influence speakers' choices of refereeing expressions (Gann & Barr, this issue), and the influence of modality and task on sentence production (Kaschak, Kutta, & Coyle, this issue). Furthermore, in keeping with the ethos of this project, as part of this introduction, we present our own review of the evidence regarding the pervasiveness of parallelism throughout the production system.

We begin with our own contribution, which aims to embody the goals of the workshop and this associated special issue by reviewing the literature with a fresh perspective. Specifically, we examine the concepts of seriality and parallelism at various levels of processing, with the aim of identifying any generalisations or common characteristics that might underpin a robust model of language production. Our survey, while not exhaustive, progresses through the stages of single word retrieval, namely lexical selection, phonological encoding, grammatical feature retrieval and syllable retrieval before moving onto sentence production. Surprisingly, with the exception of lexical selection and phonological encoding, this familiar topic, which has dominated so many other aspects of psycholinguistics theory over the decade, has received little explicit consideration in many areas of production. Following our investigation into the evidence for parallelism within speech production models, we provide a brief introduction to the other articles in this special issue.

*Corresponding author. Email: a.melinger@dundee.ac.uk
Submitted to this Special Issue following on from the International Workshop on Language Production 2010

1. Parallel processing in language production

Theories of language production propose that the speech planning process unfolds across a series of stages. Most theories posit distinct stages for at least semantic/conceptual, syntactic/grammatical and phonological/phonetic processes. Speakers must construct the meaning, grammar and sounds of individual words, and combine those words into complex expressions such as sentences. But to do this, do people construct a single representation at each stage, or do they consider alternatives in parallel, for example considering more than one candidate word at a particular point in a sentence, or more than one sound at a particular point in a word?

A great deal of research into the production of single words has been concerned with this question, specifically in terms of how activation of information spreads between lexico-semantic and morphophonological stages of processing. Different models have proposed that activation spread is interactive and continuous between levels (Dell, 1986; Harley, 1993), top-down and continuous (Caramazza, 1997; Starreveld & La Heij, 1996), or top-down and discrete (Butterworth, 1982; Fay & Cutler, 1977; Garrett, 1975, 1976, 1982; Levelt, 1989, 2001; Levelt, Roelofs, & Meyer, 1999). Continuous transmission of activation from one level to another implies that multiple alternatives must be activated in parallel at subsequent stages, whereas discrete transmission of activation is consistent with both co-activation of multiple alternatives and activation of a single candidate. As such, there has been a substantial debate about whether speakers activate a single morphophonological representation or many morphophonological representations in parallel.

The same question can of course be asked for other stages of processing, such as the phonetic or syntactic stages of processing. Yet there has been strikingly little consideration of whether the production system as a whole involves the parallel activation of alternatives. Even models of single word production that make strong claims against parallel activation during morphophonological processing have not necessarily made the same assumptions about processing at other levels. Theories of sentence production have similarly given little explicit consideration to whether alternative representations are co-active or whether processing is entirely serial (i.e., only one alternative is considered at a time).

In this paper, we review the evidence for the co-activation of alternatives at different sub-lexical, lexical and syntactic levels of representation. We also consider the evidence for different types of parallelism, a distinction of particular relevance for sentence production. Our focus is on the evidence for the co-activation and construction of non-target alternatives; we do not contrast the range of models and mechanisms proposed for selecting among candidates. We will argue that co-activation is a persistent and common feature of the production system that is relevant – albeit in different ways and perhaps to different extents – for all levels and all processes.

We first examine how serial or parallel processing could in principle be realised for lexical and syntactic processes. We then turn to relevant experimental and non-experimental evidence for parallelism. We begin with stages of single word production, focusing on activation of non-target representations at the lexico-semantic (lemma), phonological (word form) and syllabic levels of processing. We also discuss co-activation of multiple grammatical features associated with active lexical candidates, which leads in turn to the discussion of parallel processing of larger grammatically structured units, namely sentences.

2. Serial processing, parallelism of activation and parallelism of structure-building

We begin by considering the possible ways in which processing might involve the consideration of single or multiple alternatives. We propose that production processes might involve serial processing, parallelism of activation or parallelism of structure-building. The distinction between parallelism of activation and parallelism of structure building is particularly relevant for syntactic processing, but may also be relevant to those stages of lexical processing involving structure building. Note that we are interested in production-internal parallelism (i.e., activation of multiple candidates by representations at earlier stages of processing); we are not interested in spurious activation occurring through other channels (e.g., the visual or comprehension systems), although these are often exploited in experiments testing system-internal parallelism.

Under serial processing, speakers consider only a single alternative. This alternative would ideally be the intended target representation. For example, in the context of lexical processing, a speaker might activate only the single word form representation that corresponds to the representation selected at the prior stage of processing. In the context of syntactic processing, speakers might always initially choose the most frequent structure that corresponds to their intended message, and only consider any alternatives in the case of processing difficulty (i.e., during revision).

Under parallelism of activation, multiple alternative representations or structures would be activated, but selection and structure-building processes would be restricted to a single candidate. For lexical processes, we can distinguish the activation of word forms from the online construction of the phonological word. If multiple alternative word forms are activated, but only one is selected for structure building processes, we would classify this as parallelism of activation.[1] For syntactic

processing, the processor would consult the grammar to determine what syntactic possibilities were licensed under current conditions and then select which one of these alternatives to build. For example, a rule might have to reach a threshold to 'fire', and at that point, alternatives would be dropped. Thus, one rule would 'win the race' and the syntactic representation corresponding to this rule would be built; this is similar to 'race models' in the sentence comprehension literature (e.g., Frazier, 1987; Van Gompel, Pickering, & Traxler, 2001). Parallelism of activation therefore involves consideration of possible alternatives, but construction of only one alternative.

Under parallelism of structure-building, in contrast, multiple structures would be built in parallel, and some decision mechanism would be required to select between available alternatives or to stop pursuing certain options when the evidence no longer supported it as a viable candidate. For lexical processing, the speaker would begin preparing any word form receiving activation from the lexico-semantic stage. These different word forms could either be fully processed, at which point some decision mechanism would need to select the target output and discard the non-target outputs, or they could be processed only partially, eventually being discarded when their activation level fell below some threshold (see Starrevelt & La Heij, 1996, for a proposal involving a decision mechanism and Dell, 1986, for a proposal using an activation threshold). For syntactic processing, the speaker would consult the grammar to determine what syntactic possibilities were licensed under the current conditions and then build all (or many) of the alternative structures; she might immediately choose between these structures (which would equate to 'momentary parallelism' in comprehension; e.g., Altmann & Steedman, 1988), or initially pursue all of them, with all but one alternative being dropped only at some subsequent point (equating to 'extended parallelism' in comprehension; e.g., McRae, Spivey-Knowlton, & Tanenhaus, 1998).

We can apply these distinctions to each stage of processing. For lexical processing, the possibilities for parallelism of structure-building are likely limited. Lemmas, grammatical features and (at least high frequency) syllables are conceived of as unitary representations that are retrieved rather than constructed; therefore, the question of parallelism of structure-building may not be relevant. Word forms, in contrast, are both retrieved and constructed (Levelt et al., 1999). Lemmas and associated grammatical features can activate multiple morphemes which are assembled into word forms. Word forms store abstract phonological segments that need to be parsed into syllables in a context-sensitive online process. Hence, at the form level, we can evaluate the extent of parallelism for morpheme and segment assembly.

Similarly, in many accounts, syntactic processing is assumed to involve two stages: an initial stage of functional processing in which grammatical functions are assigned to unordered elements, and a subsequent stage of positional processing in which constituent structure is determined and elements are ordered (though see Cai, Pickering, & Branigan, 2012, for a one-stage model). In some models, there is a further distinction between the determination of hierarchical relations and the determination of linear order (e.g., Kempen & Hoenkamp, 1987).

In principle, then, each of these levels of syntactic processing could show parallelism of activation or parallelism of structure-building. For example, the assignment of grammatical functions might involve consultation of the grammar for licensed alternatives (agent = Subject, patient = DirectObject vs. agent = ObliqueObject, Patient = Subject) before choosing one alternative (parallelism of activation) or building functional structure associated with both alternatives (parallelism of structure-building), with one (momentary parallelism) or both (extended parallelism) being pursued. Similarly, during constituent structure processing, a speaker who wishes to describe a dative event might activate both the prepositional object rule (e.g., VP → V NP PP) and the double object rule (VP → V NP NP), but select the former to build (and not pursue the latter; parallelism of activation). Alternatively, the speaker might activate both the prepositional object rule and the double object rule, and build both of the corresponding syntactic representations (parallelism of structure-building), either immediately choosing between these structures (momentary parallelism) or initially pursuing both (extended parallelism). If linearisation of word order comprises a separate, subsequent stage of processing, multiple alternative word orders might similarly be considered and potentially pursued.

The question of whether parallel processing (and if so, of what kind) takes place at one stage of processing is of course logically independent of whether parallel processing also occurs at another stage. For example, with respect to syntax, functional vs. constituent structure processing might show different characteristics (e.g., speakers might consider multiple alternative functional assignments, but only one constituent structure). Similarly, some stages of lexical processing (e.g., lemma retrieval) might involve co-activation of multiple representations, whereas others (e.g., word form retrieval) might involve activation of only one candidate (and this is in fact the assumption made by Levelt et al., 1999).

There are some potential advantages to parallel processing of phonological and syntactic structures. Parallel phonological processing affords faster, more flexible processing as candidate representations are preactivated and potentially even partially pre-compiled (Dell, Burger, & Svec, 1997). The advantages for syntactic processing are potentially even greater. One of the strongest constraints on language production appears to be incremental processing (e.g., Levelt, 1989), so that speakers do not

plan their entire utterance in advance but instead process it in smaller units (though the scope of such incrementality remains under debate; e.g., Bock, Irwin, Davidson, & Levelt, 2003; Ferreira, 1996; Ferreira & Swets, 2002; Griffin, 2001; Griffin & Bock, 2000; Kempen & Hoenkamp, 1987; Konopka, 2012; Lindsley, 1975; Meyer, 1996; Smith & Wheeldon, 1999). Incrementality facilitates fluent production by allowing speakers to initiate speech onset without having necessarily retrieved all of the lexical content or decided the detailed syntactic structure of their intended utterance. Within such a framework, non-competitive parallel processing of syntactic structures (i.e., a race account) would maximise fluency by allowing speakers to exploit alternative structures to facilitate the integration of lexical content. For example, when a speaker describes a dative event, the lexical representation associated with the theme might be more easily accessible than the lexical representation associated with the recipient, or vice versa. If speakers were able to consider structural alternatives in parallel, they could choose whichever structure allows the more available representation to be integrated first. Non-competitive parallel processing would thus allow speakers to fully exploit the flexibility offered by the grammar in order to maintain fluency, something that would not be possible if processing was strictly serial.

On the other hand, the extent to which alternatives are considered and pursued at each level has potentially profound computational implications for processing. Although parallel activation may be cost-free, we assume parallel structure-building carries a computational burden and this should be exacerbated if more than one structure is actually pursued. For example, it is possible that extended parallelism of structure-building during both functional processing and constituent structure would involve the maintenance of huge amounts of complex and interdependent information (as Bock, 1987, noted). In the case of dative structures, for example, speakers might end up maintaining information specifying the functional assignments associated with both active and passive structures, and – for each of these assignments – both PO and DO constituent structures (and potentially both canonical and shifted PO word orders, e.g., *Buffy gave to Spike the enchanted sword*).

We might therefore expect that there would be some upper limit on the extent to which parallel structures were pursued. Furthermore, the benefits for fluent production hold only for non-competitive parallel processing: If structural alternatives compete for selection, then the availability of alternatives would tend to slow processing and increase disfluency because of the competition between them, in addition to the computational burden of maintaining multiple syntactic relations.

We now turn to the evidence for parallelism and, when appropriate, types of parallelism, at various levels of representation and processing, beginning with processing of single words and then moving on to syntactic units. Our investigation will draw on evidence from three main sources: speech errors, reaction times and choice of response type. Speech errors, especially naturally occurring errors, have the advantage of ecological validity. At the same time, they are relatively rare and have been criticised as only reflecting error-prone processes (cf. Levelt et al., 1999; Meyer, 1992). In other words, some have criticised speech errors as being uninformative about how error-free utterances are produced. Much of the reaction-time evidence comes from the picture-word interference paradigm (Rosinski, 1977). This paradigm too has been criticised for its complexity, involving production and comprehension processes. Evidence from choice of response types (e.g., choice of syntactic structure) typically provides little or no direct evidence about the timecourse of processing. Given that any one method will necessarily have limitations, we endeavour to provide evidence for parallelism from multiple sources, not relying exclusively on any single data type.

3. Lexical level

3.1. Lemma level

Lexical processing begins with the activation of lexical candidates. Most models assume that lexical processing proceeds in two stages. The first involves the selection of a *lemma*, an amodal lexical representation providing access to the word's grammatical information (e.g., Dell, 1986; Howard, Nickels, Coltheart, & Cole-Virtue, 2006; Garrett, 1984, 1992; Kempen & Harbusch, 2004; Levelt, et al, 1999; Roelofs, 1992). These lemma representations stand between the conceptual and word form levels of representation and in many models rate limit the flow of activation to later stages. Based on the top-down activation from the conceptual level, where the message, or intended meaning, has been specified, multiple lexical candidates are activated and one must be selected as the intended word. Most models allow for the free flow of bi-directional activation between the conceptual and lexical levels (but see Bloem & La Heij, 2003; Bloem, van den Boogaard, & La Heij, 2004 for an alternative model that restricts the flow of activation from the conceptual to the lexical level). Hence, almost all models assume the co-activation of lexical alternatives – that is, parallel activation of lexico-semantic representations. The basis for this key and fundamental assumption comes from a wide set of sources, including speech errors and reaction times.

Semantic substitutions, when the intended word is replaced with a semantically related intruder (e.g., *give me a fork*, intended *spoon*, from Harley, 1993), and semantic blends, when two semantically related alternatives are combined and produced as a single word (e.g., *The chung of today*, blend of *children* and *young*, from Harley, 1993),

provide strong evidence for the activation of non-target semantic alternatives. For an error of this type to occur, some unintended representations must be available for selection. These semantic errors are some of the most common types of speech error (Fay & Cutler, 1977; Garnham, Shillcock, Brown, Mill, & Cutler, 1981; Garrett, 1984) and often involve categorically related alternatives. Experimentally induced semantic substitutions demonstrated that within-category substitutions occur much more often than between-category substitutions (Vigliocco, Vinson, Indefrey, Levelt, & Hellwig, 2004; Vigliocco, Vinson, & Paganelli, 2004). Patients with semantic anomia (as a result of dementia) also reveal the activation of semantic alternatives. Most of the naming errors produced by such patients involve the production of semantic relatives of the target (e.g., Hodges, Salmon, & Butters, 1992). To account for semantic errors, Dell (1986, see also Harley, 1984) argued for the interactive spread of activation between conceptual, lemma and word form levels of representation. When trying to utter a word such as *cow*, semantically related lemmas will also be active due to spreading activation at the conceptual level. After a certain time, the most active lemma is selected. Since activation levels are dynamic and constantly changing, in some instances, the most active lemma will not be the target, and so a selection error occurs. Crucially, for present purposes, semantic errors provide evidence for the co-activation of multiple lexical candidates during lexical selection.

Errors of selection do not only occur within semantic fields. Harley (1984) noted that message-irrelevant concepts can also intrude into an utterance (e.g., intended: *I've read all my library books*, > error: *I've eaten all my library books*, in the case where the speaker reported being hungry), creating what he calls 'cognitive intrusions'. These errors illustrate that concepts activated via other cognitive systems such as perception (when the name of an event or object in the environment intrudes) or internal cognitive processes (e.g., hunger, sleep, memories) can also be co-active at the lemma level and disrupt lexical selection.

Further evidence for parallelism of activation comes from a range of reaction-time experiments. For example, classic observations from the picture-word-interference paradigm reveal slower naming times when distracting stimuli are categorically related to the target word (Glaser & Düengelhoff, 1984; Levelt, et al., 1991; Lupker, 1979; Rosinski, 1977; Schriefers, Meyer, & Levelt, 1990, Starreveld & La Heij, 1995, 1996; Underwood, 1976). Semantically related distractor words also slow down word translation, when a target word presented in one language must be translated into another (Bloem & La Heij, 2003). The standard interpretation of these effects assumes that the distracting stimulus receives converging activation both externally from the distractor word and internally from the conceptual level. When the two streams of activation converge on a non-target representation, target selection times are delayed relative to when the distractor is unrelated. This interpretation relies on the distractor's lexical representation receiving top-down activation from the conceptual system.

While the above interpretation is widely accepted, recent alternative explanations have gained some attention. The Response Exclusion Hypothesis (Mahon et al., 2007, see Dhooge & Hartsuiker, 2010, 2011, 2012 for a variant of this proposal) proposed that semantic interference effects observed in the picture-word interference paradigm arise post-lexically, not lexically. Specifically, proponents of this view argue that distractor words gain automatic access to the response buffer and must be cleared before the target response can be uttered. The time required to clear the response buffer is sensitive to factors such as category membership. Thus, categorically related distractor words slow picture-naming times, because it takes longer to clear them from the buffer than unrelated distractor words. Note that while this hypothesis does not entail the co-activation of multiple lexical candidates, proponents still assume parallel activation of lexical candidates in order to account, amongst other things, for the findings of parallel activation of multiple word forms, described in the next section.

Semantic interference is also observed with other experimental paradigms. For instance, when repeatedly naming small sets of pictures, naming times are slower (Abdel Rahman & Melinger, 2007; Belke, 2008; Belke, Meyer, & Damian, 2005; Damian, Vigliocco, & Levelt, 2001; Kroll & Stewart, 1994) and naming accuracy is lower (Schnur et al., 2006) when all the pictures are drawn from a common semantic category (a homogeneous context) than when they are not (a heterogeneous context). Crucially, the interference has been shown to extend beyond the set of presented pictures, suggesting that additional, unnamed candidates are also lexically active (Belke et al., 2005).[2]

One interpretation of the difficulty naming a set of same-category pictures is that the lexical representations of the non-target trials remain active across the naming block, inducing greater lexical competition. An alternative interpretation of this effect refers to an error-based implicit learning mechanism rather than a lexical competition mechanism (cf. Oppenheim, Dell, & Schwartz, 2010; Navarette et al., 2012). Oppenheim et al. successfully modelled the semantic interference effects computationally by adjusting the semantic-to-lexical connection weights after each trial. Crucially, both accounts assume that multiple lexical representations are activated as a result of the automatic spreading of activation between the conceptual and lexical levels. In short, while the mechanism responsible for semantic interference is under heated

debate, parallel activation of alternatives is an accepted component of lexical selection models.

3.2. Word forms

In contrast to the general consensus that non-target alternatives are active during lexical selection, the possible co-activation of phonological alternatives has been contentious. Discrete models of word production have long argued against the co-activation of phonological alternatives (e.g., Levelt et al., 1999) while non-discrete models have proposed that the active lexical candidates perpetuate their activation to subsequent levels of representation, giving rise to co-active phonological alternatives (Dell, 1986; Harley, 1993).

If multiple candidates are active at the phonological level, then evidence of their activation should be observable from naming times. For instance, if the word form of a strong competitor representation is highly active, then it should slow target naming times by direct competition either at the word form level or the lemma level, via feedback. This hypothesis has been tested extensively over the past decades and, while early attempts were unsuccessful (Levelt et al., 1991; though cf. Dell & O'Seaghdha, 1991) and therefore taken as support for discreteness, subsequent attempts have demonstrated phonologically mediated semantic interference effects.

The first demonstrations investigated pictures with two synonymous names, such as *sofa* and *couch* (Jescheniak, Hahne, Hoffmann, & Wagner, 2006; Jescheniak & Schriefers, 1998; Peterson & Savoy, 1998). Picture naming was slowed when the distractor word was related to an alternative name for the picture (e.g., preferred name *couch*, distractor *soda*, similar to the dispreferred name, *sofa;* Jescheniak & Schriefers, 1998). Similar phonologically mediated semantic interference effects have been observed for alternatives at different levels of specificity (e.g., target *fish*, distractor *card, phonologically similar to carp*; Hantsch, Jescheniak, & Schriefers, 2005), cross language translation equivalents (e.g., target *mountain*, distractor *bench, orthographically related to the Dutch word berg, which means mountain*, Hermans, Bongaerts, de Bot, & Schreuder, 1998), same level semantic alternatives (e.g., target *cat*, distractors *doll* and *log*, *both orthographically related to dog;* Abdel Rahman & Melinger, 2008)[3] and even semantic associates (e.g., target *pyramid*, distractors *camera* and *bagel*, *both orthographically related to the associate camel;* Melinger & Abdel Rahman, 2013). The distracting stimulus, which is phonologically or orthographically related to the non-target alternative, should only impact naming times if activation from the distractor word converges with top-down conceptual activation. Hence, these findings support parallelism of activation at the form level.

Further evidence for co-activation of non-target representations at the phonological level comes from the picture-picture interference task, in which participants are presented simultaneously with 2 pictures, often superimposed one over the other. Participants are asked to name one picture while ignoring the other. Target naming is faster when the names of the two pictures are phonologically related (e.g., *bottle* and *ball*) than unrelated (Morsella & Miozzo 2002; Navarete & Costa, 2005; Oppermann, Jescheniak, Schriefers, & Görges, 2010). This facilitation is consistent with converging parallel activation from the active conceptual representations of each picture onto the segments shared by both pictures. Naming times are also faster when bilinguals name objects with names that are cognates in their two languages than non-cognates (Costa, Caramazza, & Sebastian-Galles, 2000; Gollan & Acenas, 2004). Like the picture-picture facilitation effect, this implies phonological activation of the unintended name. When the target and non-target picture name share phonological content in the two languages, activation from the two word forms converge on common segmental information, speeding up retrieval and/or encoding processes.

Homophones provide another source of evidence for the co-activation of non-target word forms. Ferreira and Griffin (2003; see Cutting & Ferreira, 1999 for related findings using the picture-word interference paradigm and Burke, Locantore, Austin, & Chae, 2004, for similar findings using speed and accuracy rather than selection errors) presented participants with high-cloze sentence fragments to complete (e.g., *The woman went to the convent to become a ___ vs. I thought that there would still be some cookies left, but there were ___)*. Following each sentence trial was a picture naming trial (e.g., a picture of a priest). The sentence fragments were designed to elicit either a semantic competitor of the picture name (e.g., *nun*) or a homophone of the competitor (e.g., *none*) or matched unrelated homophonous control words (e.g., *match*). Primed sentence completion words infiltrated participants' picture naming trials, resulting in semantic substitution errors. Thus, participants accidentally named a priest as a 'nun' when the sentence fragment primed that competitor. These errors occurred both in the semantically related condition (*nun*) and in the homophone condition (*none*). This result is consistent with feedback from the word form to the lemma level, based on interactivity within the system: For *none* to affect the selection of *priest*, it must boost the activation level of *nun*'s form representation above and beyond the activation on the control word, e.g., *match*. The reinforced form representation for *nun* can then percolate back up to the lemma level to influence the selection process. Therefore, this result provides evidence for parallelism of activation at both lemma and word form levels.

Ferreira and Griffin's (2003) finding relates to the large literature on mixed errors. Although semantic and phonological substitutions are the most common type of speech error, substitutions that are both semantically and phonologically related to the intended target word occur much more frequently than would be expected by serial models (Dell & Reich, 1981; Harley, 1984; Shallice & McGill, 1978). Mixed errors have been observed in corpora of spontaneous speech errors (Dell & Reich, 1981; Harley, 1984; Harley & MacAndrew, 1995, 2001; Shallice & McGill, 1978), as well as in experimentally induced errors (Baars, Motley, & Mackay, 1975; Motley & Baars, 1976) and many aphasic individuals (Blanken, 1998; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Kulke & Blanken, 2001; Laine & Martin, 1996; Rapp & Goldrick, 2000). One common interpretation of these errors entails the activation of unintended phonological representations. Within an interactive activation model such as the one proposed by Dell (1986, see also e.g., Harley, 1984), both the intended lemma and unintended, *intruder*, lemmas spread their activation to the phonological level. When an intruder's phonology is similar to the target's, its word form activation is boosted by converging activation from the target. In models where selection takes place at the lemma level (e.g., Dell, 1986; Harley, 1993), activation from the word form level feeds back to the lemma level, thereby increasing the chances that the intruder will be selected, relative to a phonologically unrelated intruder. Hence, this interpretation, while dependent on feedback, also depends on the parallelism of activation at the word form level.

All of the evidence cited above speaks to parallelism of activation, but what evidence is there for parallelism of structure building at this stage? Word forms do not appear to be retrieved as a unit from memory but rather generated online (e.g., Dell, 1986; Dell et al., 1997; Hartley & Houghton, 1996; Levelt et al., 1999; Roelofs, 1997; Sevald & Dell, 1994; Vousden, Brown, & Harley, 2000). Unstructured segmental information is retrieved and then either parsed into larger syllabic units (e.g., Levelt, 1992) or merged with retrieved syllabic frames (Dell, 1986; 1988), a process often termed *prosodification*. Both models assume that information about stress placement is stored only when stress cannot be assigned by language-specific default rules (Schiller, Fikkert, & Levelt, 2004) and even then it is still subject to changes determined by the phonological context (Colombo, 1992; Levelt et al., 1999; Laganaro, Vacheresse, & Frauenfelder, 2002). To the extent that stress placement is spelled out or derived during phonological encoding, errors may be suggestive of parallelism of structure building at this stage of processing.

Stress placement errors are not uncommon and often emerge from conflicts between alternative forms derived from a common stem (Cutler, 1980; Fromkin, 1976), as in *The noise sort of ENvelopes you – enVElops you* and *You're in a real adVANtag-advanTAGeous position*. Here the stress pattern of the closely related alternative is suggested to intrude on target encoding. Such effects could reflect parallelism of activation (if stress assignment is lexically stored), or parallelism of structure building (if stress patterns are derived via rules or associated to syllables during prosodification; Schiller et al., 2004, but see Ferreira & Humphreys, 2001, for an alternative interpretation that does not implicate parallelism).

However, errors of stress assignment do not seem to occur across-the-board: They appear to be more common in words with irregular stress than for words with predictable rule-governed stress assignment (Cappa, Nespor, Ielasi, & Miozzo, 1997; Miceli & Caramazza, 1993; Galante, Tralli, Zuffi, & Avanzi, 2000). If parallelism of structure building consistently occurs during prosodification, we might expect that it would occur to the same extent irrespective of word type. Hence, evidence from stress assignment errors cannot conclusively support parallelism of structure building.

Perhaps, the best evidence for such parallelism comes from phonological blends involving segmental information. As discussed above, these common speech errors involve two alternative lexical options being pursued in parallel and eventually combined into a single output, e.g., *flire* (*flame* + *fire*, from Harley, 1984; see also Fromkin, 1971; Garrett, 1980; MacKay, 1982). Clearly, blending of the phonological material of two different words requires both to be pursued in parallel. However, although blends and stress assignment errors provide evidence for parallelism of structure building in error conditions, it is as yet unclear whether we have any direct evidence for parallelism of structure building in error-free conditions (i.e., parallelism might occur only occasionally, and on those occasions often lead to error).

3.3. Syllables

Many models of word production are specified only to account for semantic and phonological effects; likewise, models of articulation do not necessarily incorporate a level of phonological planning (see Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007). However, Levelt's theory includes a proposal for how phonetic syllables are retrieved after phonological encoding (Levelt, 1989, 1992; Levelt & Wheeldon, 1994). Levelt and colleagues proposed a post-phonological encoding level of representation in which precompiled syllable-sized gestural scores are stored, the *mental syllabary*. As phonological segments are combined into syllables during prosodification, all phonetic syllables in the syllabary that are consistent with the unfolding phonological syllable are activated (Roelofs, 1997). A mental syllabary at a post-phonological encoding stage is implemented in Weaver++

(Roelofs, 1992, 1997) and there is evidence consistent with it (Cholin & Levelt, 2009; Cholin, Dell, & Levelt, 2011; Cholin, Levelt, & Schiller, 2006; Laganaro & Alario, 2006). Because the model posits that most phonetic syllables are selected rather than constructed, effects at this level would again constitute an instance of parallelism of activation but not of structure building (but see Cholin, et al., 2006; Levelt & Wheeldon, 1994 who propose that low frequency and new syllables could be computed). However, to our knowledge, there is only indirect evidence, from the elicitation of speech errors, for the co-activation of multiple partially compatible phonetic-level representations.

Goldrick and Blumstein (2006) observed phonetic traces (exponents) of an intended utterance during erroneous productions of tongue twister-like utterances, such as *keff geff geff keff*, which alternated between voiced and voiceless initial stop segments. In English, word initial voiced and voiceless plosives differ not only in terms of voicing but also voice onset time (VOT), namely the asynchronous alignment of the release of the closure by the articulators (lips, tongue) and the onset of voicing. Goldrick and Blumstein measured the length of the VOT in intended and unintended segments (i.e., in errors) and found shorter VOTs in unintended than intended voiceless segments, but longer VOTs in unintended than intended voiced segments. In other words, they found acoustic traces of the target on the error, resulting in a sort of phonetic blend. They argued that the contamination of the articulatory plan provided evidence for cascading activation from phonological representations to articulatory processes, implying the activation of non-target phonetic representations (see Goldstein et al., 2007; McMillan, Corley, & Lickley, 2009; McMillan & Corley, 2010, for additional examples of 'contaminated' phonetic realisations).

These phonological 'traces' are also found in non-erroneous speech, for example from incomplete neutralisation. In these instances, a phonemic contrast is lost due to a phonological process, such as word final devoicing. Close phonetic inspection reveals that the two apparently neutralised surface forms reveal traces of the underlying phonemic content (Charles-Luce, 1985; Port & O'Dell, 1985; Slowiaczek & Dinnsen, 1985; Warner, Jongman, Sereno, & Kemps, 2004). Thus, parallel activation of alternative phonetic representations is found in error-free utterances as well. Hence, while there does not appear to be any direct investigation of whether multiple syllabic representations are activated in parallel, the indirect evidence from speech errors and incomplete neutralisations is consistent with this claim.

3.4. Syntactic features

Syntactic and morphosyntactic features associated with a lexical item can be characterised either as context-sensitive 'extrinsic' features or lexically specified 'intrinsic' features. For the former, nouns can be marked for number (singular, plural) or case (nominative, accusative, dative, etc) while verbs can be marked for tense (present, past, future), aspect (perfect, progressive, etc), person (1st, 2nd, 3rd), or number (singular, plural). The value of a particular extrinsic feature depends on the context in which it is being used. In contrast, intrinsic features are not defined by the context. For example, content words are associated with a part of speech (noun, adjective, verb), which has both morphological and syntactic implications, and in some languages with grammatical gender (feminine, masculine, neuter, common, etc.), countability, or noun classifiers. Verbs can be associated with features specifying auxiliary type (e.g., in French: *être, avoir*) and with subcategorisation information indicating their combinatorial potential (e.g., whether a verb can co-occur with a noun phrase or a prepositional phrase). Intrinsic features allow us to ask whether features associated with lexical alternatives are activated in parallel. Subcategorisation features additionally allow us to ask whether one lexical item can activate all its compatible combinatorial alternatives in parallel. We focus here on the former question and revisit the latter in Section 4.

It is generally assumed that words that share a syntactic characteristic are associated to a common syntactic or morphosyntactic feature; for instance, all nouns are linked to a common NOUN node (Caramazza, 1997; Dell, 1986; Gaskell & Marslen-Wilson, 1997; Levelt et al., 1999; Pickering & Branigan, 1998). Interestingly, despite architectural and processing differences, models generally share the assumption of parallel activation of grammatical features. This means that co-active lexical alternatives make their respective grammatical features partly available. For example, *advice* (a mass noun) and *recommendation* (a count noun) would partially activate mass and count features, respectively, during lexical selection. The assumption of parallel co-activation of grammatical features largely follows from the interpretation of other behavioural observations, rather than direct investigation.

Evidence of parallel activation of grammatical features is found in speech errors. Berg (1992; Marx, 1999, see also Vigliocco et al., 2004, for experimental demonstration) observed that in spontaneous German semantic substitutions, targets and intruders tend to have the same grammatical gender (an effect he termed the *identical gender effect*) and grammatical class (the *syntactic category constraint*; Garrett, 1975; Stemberger, 1985). Berg proposed that active lemmas must activate their associated gender features before they are selected.[4] If we accept this proposal in tandem with the well-supported claim that multiple lemmas are co-active in parallel (e.g., Levelt et al., 1999), we can conclude that all active lemmas make their grammatical features, or minimally their gender and part of speech features, active prior to selection. An alternative account

attributes the syntactic category constraint to an interaction between the syntactic and lexical processes, but similarly assumes parallel activation of grammatical features (Dell, Oppenheim, & Kittredge, 2008.)

The tip-of-the-tongue phenomenon is also informative about parallel activation of grammatical features. Speakers experiencing a tip-of-the-tongue state can often access grammatical information such as grammatical gender or mass/count features without being able to access phonology (Caramazza & Miozzo, 1997; Miozzo & Caramazza, 1997; Vigliocco, Antonini, & Garrett, 1997). For one-stage models that do not include a lemma level, such as the Independent Network Model (Caramazza, 1997), this finding entails parallelism. Caramazza suggested that grammatical features are (usually) selected before their corresponding lexical item. For lexically specified grammatical features to be available for selection before the target word is selected, the model must ensure that active lexical items activate their grammatical features in parallel.

These examples relate to cases where multiple co-active lemmas or word forms make their respective intrinsic grammatical feature specifications partially available (i.e., there is many-to-many parallelism of activation). But grammatical features serve to allow the construction of larger syntactic units, by definition a process of structure building. Furthermore, a single lemma can be compatible with more than one specification of a particular feature. For example, most verbs are associated with multiple subcategorisation features. The verb *to eat* can be associated with both the intransitive frame [NP ___], as in 'I am eating' or the transitive frame [NP ____ NP], as in 'I am eating pizza'. Therefore, we can examine the evidence for one-to-many activation as well. However, the activation of subcategorisation frames is best assessed in the context of sentences, not single words.

4. Syntactic processes

The issue of how parallel activation and structure building impacts on the processing of syntactically complex units can be investigated from a number of perspectives. For instance, multiple lexical representations may be activated at the same time, although they are intended to be produced at different points in an utterance. For example, Meyer (1996) found that speakers showed interference when describing arrays involving two objects (e.g., *The arrow is next to the bag*) when they were exposed to auditory distractors that were semantically related to either object, suggesting that lexical representations for both objects were active prior to speech onset (see also Wagner, Jescheniak, & Schriefers, 2010). She also found comparable interference in phrases such as *The arrow and the bag* when a distractor was related to the second noun as to the first noun, though more recent evidence has suggested that the first noun is processed more thoroughly than the second noun (Allum & Wheeldon, 2009). Similarly, Smith and Wheeldon (2004) found that nouns from the same phrase were more likely to be processed in parallel than nouns from different phrases, such that speakers were slower to initiate sentences such as *The saw and the axe moved down* than *The saw moved below the axe*. In contrast, Griffin (2001) showed that when speakers produced sentences such as *The clock and the television are above the needle*, speech onset times varied according to the codability and frequency of the first noun alone (see also Griffin & Spieler, 2006; Meyer, Roelofs, & Levelt, 2003; Meyer, Sleiderink, & Levelt, 1998).

Note that simultaneous activation of words associated with different positions in the same utterance requires some mechanism by which speakers can ensure that the right word is produced at the right moment, that is, some means of integrating activation within the lexicon with syntactic processes. Dell et al. (2008) provided such an account, in which syntactic units (representing syntactically defined sequential states) affect the activation of lexical elements through excitatory and inhibitory connections.

The work described above is not, however, directly relevant to the question of whether people consider alternative possible realisations of individual units in production. The question of whether *arrow* and *bag* are activated together when uttering an expression containing both words is not the same as whether, for example, *sack* is activated as well as *bag* when *bag* is eventually uttered.

In this context, we can consider the evidence from subject-verb agreement. Bock and Miller (1991) found that speakers were more likely to produce plural-inflected verbs following singular subject noun phrases including a plural embedded noun (e.g., *the report of the fires were …*) than singular subject noun phrases including a singular embedded noun (e.g., *the report of the fire were …*). At least on those occasions when people make errors, they must consider the syntactic number features of *report* and *fires*. In itself, this may simply demonstrate that the grammatical features of both nouns are simultaneously activated and, thus, does not demonstrate parallelism in the relevant sense. But the finding that both *report* and *fires* can affect the number of the following verb suggests (though does not prove) that speakers may attempt to form two dependencies – subject-verb and embedded noun-verb – at the same time, in which case dependency formation would involve parallelism of activation. A large number of experiments have demonstrated that agreement can be affected by more than one noun in this way, in different languages and constructions (e.g., Vigliocco, Butterworth, & Semenza, 1995; Hartsuiker, Antón-Mendez, & Van Zee, 2001).

Reaction time data provide converging evidence in support of this account. Staub (2009, 2010) presented participants with subject phrases containing number-matching

and number-mismatching attractors, and had them make a speeded choice between singular and plural verb forms (e.g., *is* vs. *are*). Participants were slower to make correct responses when the attractor mismatched the head noun in number than when it matched. These findings suggest that people simultaneously consider alternative dependencies in a way that affects response times, and that these effects are not limited to a small proportion of potentially erroneous responses.

Taken together, these results suggest that more than one lexical representation may be simultaneously available during syntactic processing, and moreover may interfere with the accuracy of syntactic processes. However, they are not informative about another important question: whether language production involves simultaneous activation of more than one syntactic representation, in other words whether speakers simultaneously consider alternative syntactic realisations of a message. In fact, as we noted in the introduction, this question has received little attention in production research. Research has often made implicit assumptions about the parallel activation of syntax, but there has been little explicit consideration of the extent to which production involves consideration of multiple possible syntactic structures. This omission is particularly striking given the centrality of questions about the parallel versus serial nature of syntactic processing in theories of language comprehension (e.g., Frazier, 1987; MacDonald, Pearlmutter, & Seidenberg, 1994). Here, we consider three types of evidence that might adjudicate between the alternative possibilities: evidence from speech errors, evidence from the timecourse of processing, and evidence from the syntactic choices that speakers make.

4.1. Syntactic blend errors

As with other levels of structure, speech errors are potentially informative about the extent to which speakers consider more than one syntactic realisation for their message. As noted by Bock (1987), the occurrence of syntactic blend errors suggests that the processor activates and builds structure associated with more than one syntactic rule (or sets of rules) under at least some circumstances, and that these structures may be processed to the point of articulation. In such errors, the speaker appears to blend two alternative syntactic formulations of (roughly) the same message, for example *This is getting very difficult to cut this* (intended: *It is getting very difficult to cut this/This is getting very difficult to cut;* Stemberger, 1985), *Would you turn on the light on* (intended: *Would you turn the light on/Would you turn on the light*; Fay, 1980), *This is so much exciting!* (intended: *This is so much fun/This is so exciting;* author's own error). Coppock (2010) found that the distribution of a set of blend errors in a corpus was consistent with the effect having a syntactic rather than an articulatory locus (i.e., being associated with grammatical encoding rather than articulation): When other factors (e.g., prosody) were controlled for, genuine errors were more likely to involve the overlap of syntactic features between the two elements comprising the blend, relative to a set of randomly generated, unattested blend errors.

Syntactic blend errors suggest co-activation of syntactic alternatives, but they are compatible with more than one type of co-activation. They could occur, because the message leads to activation of two syntactic rules; rather than one rule being selected, the processor selects a blend of the two rules, and uses this blended rule to construct a blended structure that is partially compatible with each of the correct rules (parallelism of activation). Alternatively, the processor might select both rules and construct two distinct structures, which are then blended, presumably to construct a single structure (parallelism of structure-building). In that case, the fact that the utterance that the speaker ultimately produced involved elements of both structures would constitute evidence for extended parallelism.

4.2. Timecourse of processing

The time taken for speakers to initiate utterances in the absence or presence of syntactic alternatives is also potentially informative about the extent to which they consider more than one syntactic option. If syntactic processing is serial, then it should make no difference whether there is more than one alternative structure available: The speaker should choose the preferred structure on the basis of whatever criterion is usually used, and ignore the alternative possibility. Thus, response times should be the same whether only one or more than one syntactic structure is possible. In contrast, if syntactic processing takes place in parallel, the existence of alternatives should affect processing times. If alternative structures compete for selection, processing should be slowed (and more error-prone) in the presence of alternatives, because competition between alternatives must be resolved; if alternative structures do not compete (i.e., a race model), then processing should be faster in the presence of alternatives, because this allows the accommodation of variations in lexical accessibility by allowing early production of the more highly available item.

V. Ferreira (1996) tested these possibilities using a paradigm in which speakers were given a set of words and asked to produce an utterance containing those words. He manipulated the availability of one or more than one syntactic structure by using verbs that allowed both PO and DO structures (e.g., *give*) or only PO structures (e.g., *donate*), and by requiring participants to use the preposition *to*. Participants were faster (and less error-prone) to respond in conditions that allowed syntactic flexibility, in other words when they produced sentences involving verbs that allowed both PO and DO structures, and were

not required to use *to*. Ferreira argued that these results support non-competitive parallel processing of syntactic structure, in which more than one structure is simultaneously available for integration with lexical content. Assuming that PO and DO structures differ only in their constituent structures (i.e., do not differ in their functional structures), then this would suggest that constituent structures are considered in parallel.[5]

Other evidence about the timecourse of processing also supports the assumption of parallel processing, but argues against a parallel race model. Stallings, MacDonald, and O'Seaghdha (1998) used a paradigm similar to that used by Ferreira (1996) to examine the production of PO and shifted PO sentences (e.g., *Janet revealed some more specific plans for a brand new defense plant to Leigh* vs. *Janet revealed to Leigh some more specific plans for a brand new defense plant*), using verbs that were commonly associated with an adjacent NP (e.g., *transfer*) or a non-adjacent NP (e.g., *reveal*). When producing sentences involving verbs that were commonly associated with non-adjacent NPs, participants' decision latencies were slower when producing the standard order and faster when producing the shifted order. Stallings et al. argued that these results are compatible with parallel competitive processing, in which alternative analyses were simultaneously active and competed with each other for selection. Given that PO and shifted PO structures have the same functional structure (and so would not compete at this level of structure), these results suggest that parallel competitive processing occurs at the level of constituent structure.

Further support for a competitive model comes from Hwang and Kaiser (in press), who used the same paradigm as Ferreira (1996) to investigate the production of numeral-quantifier constructions and active/passive constructions in Korean. They found that speakers had longer onset latencies when their utterances could also be expressed a different syntactic construction than when they could not (see also Myachykov, Scheepers, Garrod, Thompson, & Fedorova, 2013).

Thus evidence from the timecourse of processing suggests parallel processing of some aspects of syntactic structure, but does not determine the precise nature of that parallelism. In part, the contradiction between existing findings may reflect the difficulty, when analysing response latency data, in isolating factors associated with structural processing from those associated with other aspects of processing such as lexical retrieval (see e.g., Konopka, 2012).

4.3. Evidence from speakers' syntactic choices

A number of studies examining the factors that determine speakers' choices between structural alternatives are also compatible with, and in some cases strongly supportive of, the co-activation of syntax. We have noted that there is generally more than one way to realise a message with respect to syntactic structure, and speakers must therefore have some way of deciding between alternatives. Many studies have shown that speakers' default preferences for particular alternatives are not fixed and can be altered, depending on linguistic and extralinguistic factors.

This tendency towards flexibility of syntactic choices is difficult to reconcile with serial processing, which would predict speakers to make consistent choices. If speakers used only one factor to make their decisions, it would be quite straightforward to consider the alternatives in serial or in parallel. For example, there is a strong overall preference in English for active over passive sentences (as they are more frequent, and perhaps less complex). If speakers needed to choose between an active and a passive realisation of a transitive event, they could consider the more frequent active first, then abandon it if it was problematic (e.g., if it were inappropriate or there was difficulty in retrieving the agent) and instead select the passive realisation. Alternatively, they could consider both structures together, in order to make their decision.

But the serial account becomes much more difficult to sustain when there is more than one relevant constraint. Thus, in some circumstances, the overall preference for actives may play off against another consideration that would favour a passive, as we discuss below. It is easy to accommodate more than one constraint within a parallel account: Speakers can weigh up the strength of evidence for each alternative structure and select among them (either immediately or after some interval). If the speaker considered the alternatives in serial, she would have to select one using one constraint (e.g., active), then determine if it was sufficiently compatible with the other constraint, and reanalyse if it were not. This is unwieldy and inefficient when there is more than one constraint, and becomes increasingly so with the addition of further constraints, potentially requiring multiple decisions to be made and then revised. It is easy to show that it will not lead to all constraints being given an appropriate weighting. Thus, evidence that speakers are sensitive to more than one constraint in making their syntactic choices would provide evidence to support parallel processing, with this evidence being stronger for demonstrations of multiple constraints.

Many studies have shown that speakers' syntactic choices are affected by factors related to the accessibility of verb arguments. Bock and Warren (1985) showed that speakers tended to choose structures that allowed more concrete entities to appear in higher grammatical functions. For example, participants were more likely to erroneously recall passive sentences as actives when the agent was more concrete and the patient was less concrete (e.g., recalling *The shock was administered by the doctor* as *The doctor administered the shock*) than vice versa. A

large number of studies of English and other languages have subsequently found that speakers' choices between active and passive structures are similarly affected by variations in animacy, so that they choose structures that allow entities higher up the animacy hierarchy (e.g., humans) to appear as sentence subjects (e.g., McDonald, Bock, & Kelly, 1993; Prat-Sala & Branigan, 2000; Tanaka, Branigan, McLean, & Pickering, 2011; van Nice & Dietrich, 2003). Other studies in English and Spanish have shown that the discourse status of arguments affects syntactic choices, with speakers tending to produce structures in which given or contextually more salient entities appear as the subject (e.g., Bock, 1977; Bock & Irwin, 1980; Gennari, Mirkovic, & MacDonald, 2012; Prat-Sala & Branigan, 2000; Turner & Rommetveit, 1967).

In all of these studies, ultimate structure choice was determined by an interaction of default structural preferences and a factor that affected the relative accessibility of arguments. These results suggest that choice of functional assignment (agent = Subject, patient = DirectObject vs. agent = ObliqueObject, patient = Subject) is susceptible to more than one influence, and hence support parallel processing at the level of functional processing. However, they do not determine whether such parallelism is associated with activation or building of alternative functional structures, and if the latter, whether parallel processing is momentary or extended. Nor are they informative about whether parallelism is competitive or non-competitive.

Other studies have shown that the same factors may also influence speakers' choices regarding aspects of syntactic structure other than grammatical function. In Odawa, animacy affects choice of active or inverse structure (which differ in thematic emphasis), when both alternatives have the same grammatical function assignment (Christianson & Ferreira, 2005). In other languages, they affect word order. For example, Greek, Spanish and Japanese all allow the object to precede the subject as well as vice versa, though subject–object order tends to be strongly preferred by default. Experiments using a sentence recall paradigm have shown that in all three languages, the overall preference to produce subject-order is mediated by animacy, with speakers more likely to produce object–subject order when this allows an animate entity to precede an inanimate entity (Branigan & Feleki, 1999; Prat-Sala & Branigan, 2000; Tanaka et al., 2011; see also Gennari et al., 2012). For example, Greek participants tended to recall *Sta dimokratika politevmata, to sindagm sevete ton politi*; 'In democratic regimes, the law$_{SUBJ}$ respects the citizen$_{OBJ}$' as *Sta dimokratika politevmata, ton politi sevete to sindagm*; 'In democratic regimes, the citizen$_{OBJ}$ respects the law$_{SUBJ}$ (Branigan & Feleki, 1999). Similarly, givenness and discourse salience interact with word-order preferences, so that speakers are more likely to produce object–subject order when this encodes given or salient information before new or non-salient information than vice versa (Ferreira & Yoshita, 2003; Prat-Sala & Branigan, 2000).

A further factor, constituent weight, has also been shown to affect constituent structure as well as word order. In both Japanese and English, speakers override canonical word order preferences in response to constituents that are particularly 'heavy'. In Japanese, speakers' likelihood of producing disfavoured OSV order for transitive events increases when the object NP is heavy, yielding a long-before-short preference (Yamashita & Chang, 2001). In English, speakers' likelihood of producing the strongly disfavoured V-PP-NP order for dative events increases when the theme NP is relatively heavier than the beneficiary, giving rise to a short-before-long preference (Arnold, Losongco, Wasow, & Grinstrom, 2000, Wasow, 1997a, b). But Arnold et al. also found that constituent weight affected choice of constituent structure, with more beneficiary-early constructions when the theme was heavier than the beneficiary.

Other results suggest that default preferences for particular word orders or constituent structures are mediated by recent structural experience. Bock (1986) showed that speakers' default preference for active structures is reduced if they have recently experienced passive structures, and similarly for PO and DO structures. She argued that this syntactic priming effect is associated with alterations in the accessibility of particular constituent structure representations (Bock, 1989; Bock & Loebell, 1990). Subsequent studies have demonstrated similar effects for other alternations, involving choices between both constituent structure (e.g., the structure of complex noun phrases; Cleland & Pickering, 2003) and word order (e.g., the order of main verb and auxiliary in Dutch; Hartsuiker & Westenberg, 2000).

Overall, these results suggest that – like functional structure – constituent structure and word order are affected by more than one factor. As such, they appear to be more straightforwardly explained in terms of parallel than serial processing. Moreover, they provide some evidence that such parallel processing is not restricted to simple activation of available rules or momentary structure-building. The fact that effects of animacy manifested themselves in both choice of a passive structure (assigning an animate patient to the subject function) and choice of word order (assigning an animate entity to first position in the sentence) suggests that speakers did not choose between an active and a passive structure during functional processing, but instead pursued both possibilities through to positional processing (i.e., parallelism of structure building during functional processing), yielding both passive and non-canonical (OSV) active responses (Tanaka et al., 2011). Similarly, if syntactic priming effects operate over constituent structure, then the

existence of syntactic priming for active versus passive structures implies that speakers must have built and pursued two different grammatical function assignments, one associated with an active structure and one associated with a passive structure, with syntactic priming determining which functional assignment is successfully realised as a constituent structure (Bock, 1986). If the speaker had not pursued both functional structures, there would be no opportunity for syntactic priming to come into play (because the chosen functional structure would mandate which constituent structure must be built).

However, even stronger evidence to support parallel processing of syntactic structure comes from a series of demonstrations that syntactic choice is reliant upon an interaction of multiple factors. In such cases, it is very difficult to see how serial processing could occur without requiring considerable reanalysis and disfluency. F. Ferreira (1994) showed that choice of active versus passive structure was simultaneously influenced by both verb type and animacy: When participants were asked to produce English sentences on the basis of two nouns and a verb, they were more likely to produce passive sentences when there was a mismatch in animacy, but this tendency was more marked for verbs with theme-experiencer thematic role structure (e.g., *annoy*) than for verbs with agent-theme or experiencer-theme structure (e.g., *hit, hate*). Gennari and MacDonald (2009) found a similar pattern in the production of English relative clauses. When they provided participants with words to form into a complex expression, using theme-experiencer and experiencer-theme verbs (e.g., *writer, critic/review, angered/analysed*), participants were more likely to produce passive relative clauses with theme-experiencer verbs when the two nouns differed in animacy than when they were both animate; in the latter case, speakers produced passives at the same rate as with experiencer-theme verbs (when both nouns were animate). Prat-Sala and Branigan (2000) showed a similar relationship between discourse salience and animacy: The tendency for speakers to produce English and Spanish passive sentences when the patient was salient was stronger when it was also animate than when it was inanimate.

Several studies have demonstrated multiple determinants of constituent structure, with most studies focusing on word order. Prat-Sala and Branigan (2000) demonstrated that discourse salience and animacy also interacted in Spanish speakers' choice of canonical versus left-dislocated active structures, with speakers most likely to produce left-dislocated structures (e.g., *A la mujerla atropello el tren* 'The woman, the train ran her over') when the patient was both salient and animate. Stallings et al. (1998) found that choice of canonical versus shifted PO structures was affected by both constituent weight and verb-specific preferences, so that shifting was more likely when the theme was heavy and the verb was one that more frequently appeared with a non-adjacent NP. Arnold et al. (2000) similarly showed in a corpus study that the choice of canonical versus shifted structures was associated with both givenness and syntactic weight: Shifting was most likely when the theme was both new and heavy. They found the same pattern of given-before-new and light-before-heavy in an experimental study of the dative alternation. However, their results showed that the role that each factor played depended in part on the strength of competing factors. Discourse status played a relatively weak role when constituents differed greatly in syntactic weight; in contrast, when one argument was given (and the other was new), discourse status exerted a stronger influence than syntactic weight. They suggested that this pattern is consistent with a constraint-based production system, where the strength of one constraint varies in response to the strength of competing constraints – in other words, competitive parallelism of structure building (with respect to constituent structure). However, the data do not determine whether such processing involves parallel activation or parallel structure building.

Gennari et al. (2012) provided further evidence for competitive parallelism of structure building of both functional assignment and constituent structure. In their study, English, Spanish and Serbian speakers produced relative clauses describing events involving an animate agent and an animate or inanimate patient. In all three languages, animacy affected the choice of active or passive structures (though to differing degrees), but also whether the agent was overtly expressed or not, with speakers being more likely to omit (animate) agents when the patient was animate than inanimate (e.g., agentless passives: *The man being punched*; or impersonal actives: Spanish *El hombre (al) que están golpeando* 'The man being punched'). Moreover, the likelihood of omitting the agent correlated with the semantic similarity of animate agent and patient entities. They argued that speakers simultaneously consider both entities during syntactic processing and experience more competition if the entities are semantically related; this leads to inhibition of the linguistic representation for the unselected entity and hence its omission. In other words, they argued that sentence production involves competitive parallel activation during both functional and constituent-structure processing.

In summary, unlike lexical processing, there has been little direct debate about the existence or extent of parallel processing in sentence production. However, many lines of evidence suggest that speakers consider alternative syntactic representations in parallel, with respect to the formation of agreement dependencies, grammatical function assignment and the construction of constituent structure and word order. At least, some of this evidence is consistent with not just parallel activation of

alternatives, but also (possibly extended) parallel structure building.

5. Conclusions

The question of whether speakers consider more than one alternative in parallel before eventually making their choice about what to say has received a great deal of explicit investigation with respect to some linguistic representations, and much less with respect to others. If it were the case that speakers do this, and in the same way, at every level of processing, then although sounds, words and sentences involve different types of representations, they would share an important aspect of their underlying mechanisms and the production system would operate in a consistent manner. However, consistency across the production system need not be the case. As our review demonstrates, there is evidence for parallelism at many stages of production. We have argued for two distinct types of parallelism, namely parallelism of activation and parallelism of structure building. At the lexical level, evidence for the former is strong. There is less evidence for the latter, in part – we have suggested – because much lexical processing involves retrieval rather than structure building. At the syntactic level, however, there is relatively strong evidence for parallelism of both activation and structure building, although it is not yet clear whether the production system builds parallel structures whenever possible, or only in some cases (e.g., for certain aspects of syntactic structure).

Overall, the evidence for parallelism in production is stronger at some levels than at others and it is premature to make a general claim about its ubiquity. But it is a plausible hypothesis that the language production system is fundamentally parallel in its operation, and we believe that testing it serves as a clear and fundamental goal for researchers who focus on specific levels or stages of production.

Our survey does not of course address all aspects of production. On the one hand, we have said nothing about the process of articulation, whether people activate motor programmes in parallel when planning movements of the tongue or lips. At the other extreme, we have not considered whether speakers consider alternative plans of what to talk about, or indeed other aspects of what Levelt (1989) refers to as ‘macroplanning’.

Another obvious issue for future investigation constitutes parallelism in bilinguals. It is theoretically possible that a bilingual’s two languages are separated so that the use of one language does not lead to the activation of the other language. However, there is extensive evidence that bilinguals regularly activate aspects of the non-target as well as the target language (see Kroll, Bobb, Misra, & Guo, 2008, for a review). It is therefore plausible that parallel activation occurs between languages as well as within a language, again at the different levels of representation concerned, for instance, with sounds, words and syntax. Further questions relate to the extent of parallel activation in children, older adults, and people with language disorders such as aphasia. Presumably parallel activation is not limited to adults without language impairment. But it may be that the extent of parallelism depends on the population in question. In conclusion, there is much evidence that language production involves parallel processing of alternatives, but the extent to which parallelism is a consistent characteristic of the production system is unclear and remains an important question for future research.

6. Introduction to the special issue

Having reviewed the evidence for parallelism across different levels of processing across the production system, we now briefly introduce the other papers contained within this Special Issue. A central question in the domain of language production, which was flagged in our own review, concerns the mechanism by which a single word is selected from the lexicon and prepared for articulation. The answer might depend on the type of word or morpheme being retrieved. Content words, such as nouns and verbs, and function words, such as determiners and prepositions, have different properties and are sensitive to different types of information, supporting the possibility that they might be selected via different mechanisms. The proposal that different types of words have different selection processes has a long tradition in both the comprehension and production literatures (e.g., Garrett, 1982; Kimball, 1973, see also Bock, 1989). To support these ideas, Janssen and Caramazza (2003) outlined some of the differences in the types of information that are relevant for the selection of nouns as opposed to determiners in different languages. Pushing this idea further, Costa, Kovacic, Fedorenko, and Caramazza (2003) suggested that bound and freestanding morphemes might also be selected by different mechanisms. This possibility is taken up in the contribution by Jescheniak, Schriefres, and Lemhöfer (this issue_a). Specifically, they review the literature on how freestanding and bound gender-marking morphemes are selected, and argue for a single selection mechanism. In a reply, Janssen et al. (this issue) take issue with this interpretation of the literature and offer new data that challenge it. This discussion is continued in a rejoinder by Jescheniak et al. (this issue_b).

In their contribution, Perret et al. (this issue) examine the locus of syllable frequency effects. Behavioural studies into word form preparation have implicated a store of syllable-sized phonetic plans (Levelt & Wheeldon, 1994). These syllable representations appear to be relevant during phonetic planning but not during phonological planning (Schiller, 1998). However, evidence from

brain-damaged speakers is not totally consonant with this conclusion. Some speakers with phonological impairments still display syllable-specific deficits (e.g., Laganaro, 2005, 2008). Perret et al. (this issue) examine the source of these discrepancies by presenting a parallel investigation of syllable frequency effects with healthy and impaired speakers. They conclude that the discrepancies are not attributable to methodological differences present in the literature and are better accounted for by interactions between phonological and phonetic encoding.

Following Bock (1986), many studies have shown that people tend to repeat the syntactic structure of utterances they have produced or comprehended (see Pickering & Ferreira, 2008). Quite remarkably, such effects can last for at least a week, so long as participants are first exposed to multiple (cumulative) utterances with the same structure (Kaschak, Kutta, & Schatschneider, 2011). In their contribution, Kaschak et al. (this issue) report cumulative priming within a session and over a week when the task did not change (picture description). When the task changed from picture description to written stem completion, or vice versa, priming persisted within the session but did not persist over a week. Priming therefore transfers across tasks in the relatively short term and persists for a very long time, but such priming appears to be task specific. The authors interpret their results in terms of theories of priming and memory.

Gann and Barr (this issue) investigate the conditions under which people tailor their utterances according to their addressees' needs (i.e., use audience design; Clark & Murphy, 1982). They report an experiment in which participants repeatedly referred to both conventional and unconventional objects in an array, under conditions where expanded descriptions (e.g., *acoustic guitar*) were needed for disambiguation. They then referred to both those and new objects under conditions where a simpler description (*guitar*) would be adequate, either to the same or a new addressee, and under conditions when the addressee could or could not provide online feedback. Participants regularly overspecified descriptions but were affected by the feedback provided by their addressees. They lengthened their descriptions for new addressees (in accordance with audience design), even when their addressees did not provide feedback. The authors argue that such adaptation is not pre-planned but occurs through processes of monitoring and adjustment (e.g., Horton & Keysar, 1996).

Notes

1. Note that word forms are often composed of multiple morphemes, as a result of inflexions (e.g., *cats, walking*) and derivations (e.g., *unhappy, quickly*). While stem and derivational morphemes are activated by their corresponding lemmas, inflexions are activated by diacritic features associated with the lemma or the conceptual structure of the utterance (Levelt, 1989; Janssen, Roelofs, & Levelt, 2002). While the morphemes that make up a given word form are likely activated in parallel, this is not the type of parallelism under discussion here. Rather, we are concerned with the parallel activation of *alternatives*, such as cat vs. dog.
2. While most of the evidence for co-activation of lexical alternatives has focused on same-category alternatives, there is evidence for the co-activation of different-category semantic associates as well (Abdel Rahman & Melinger, 2007, 2011; Melinger & Abdel Rahman, 2013). Similarly, studies have demonstrated that the names of physical features of objects, such as their colour, are active during object naming (Kuipers & La Heij, 2009; Navarete & Costa, 2005). These findings demonstrate that lexical co-activation is a general property and not restricted to a single class of semantic relation.
3. Abdel Rahman and Melinger (2008) and Melinger and Abdel Rahman (2013) presented two related or unrelated distractor words together with the target picture.
4. Note that to account for the identical gender effect, Berg also argued for bi-directional links between lemmas and grammatical features.
5. It has also been suggested that they differ in their functional structures (e.g., Bock & Warren, 1985); if so, then such evidence would support the parallel processing of functional structures but would not necessarily be informative about the parallelism of constituent structure processing.

References

Abdel Rahman, R., & Melinger, A. (2008). Enhanced phonological facilitation and traces of concurrent word form activation in speech production: An object naming study with multiple distractors. *Quarterly Journal of Experimental Psychology*, *61*, 1410–1440. doi:10.1080/17470210701560724

Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*, 149–161. doi:10.1037/a0021208

Abdel Rahman, R., & Melinger, A. (2007). When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 604–614. doi:10.1037/0278-7393.33.3.604

Allum, P., & Wheeldon, L. R. (2009). Scope of lexical access in spoken sentence production: Implications for the conceptual-syntactic interface. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*, 1240–1255. doi:10.1037/a0016367

Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238. doi:10.1016/0010-0277(88)90020-0

Arnold, J. E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, *76*, 28–55.

Baars, B. J., & Motley, J. T., & MacKay, D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382–391. doi:10.1016/S0022-5371(75)80017-X

Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin and Review*, *15*, 357–363. doi:10.3758/PBR.15.2.357

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology: Section A, Human Experimental Psychology*, *58*, 667–692. doi:10.1080/02724980443000142

Berg, T. (1992). Prelexical and postlexical features in language production. *Applied Psycholinguistics*, *13*, 199–235. doi:10.1017/S0142716400005567

Blanken, G. (1998). Lexicalisation in speech production: Evidence from form-related word substitutions in aphasia. *Cognitive Neuropsychology*, *15*, 321–360. doi:10.1080/026432998381122

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, *48*, 468–488. doi:10.1016/S0749-596X(02)00503-X

Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language*, *51*, 307–323. doi:10.1016/j.jml.2004.05.001

Bock, J. K. (1977). The effect of a pragmatic presupposition on syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior*, *16*, 723–734.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387. doi:10.1016/0010-0285(86)90004-6

Bock, J. K. (1987). Coordinating words and syntax in speech plans. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 3, pp. 337–390). London: Erlbaum.

Bock, J. K. (1989) Closed-class immanence in sentence production. *Cognition*, *31*, 163–186. doi:10.1016/0010-0277(89)90022-X

Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39. doi:10.1016/0010-0277(90)90035-I

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45–93. doi:10.1016/0010-0285(91)90003-7

Bock, J. K., & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior*, *19*, 467–484. doi:10.1016/S0022-5371(80)90321-7

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*, 653–685. doi:10.1016/S0749-596X(03)00007-X

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*, 47–67. doi:10.1016/0010-0277(85)90023-X

Branigan, H. P., & Feleki, E. (1999). Conceptual accessibility and serial order in Greek language production. In M. Hahn & S. C. Stoness (Eds), *Proceedings of the 21st Conference of the Cognitive Science Society*, (pp. 96–101). Mahwah: Erlbaum.

Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults' production of proper names. *Psychological Science*, *15*, 164–170. doi:10.1111/j.0956-7976.2004.01503004.x

Butterworth, B. (1982). Speech errors: Old data in search of new theories. In A. Cutler (Ed.), *Slips of the tongue* (pp. 627–662). Amsterdam: Mouton.

Cai, Z., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language*, *66*, 833–849. doi:10.1016/j.jml.2012.03.009

Cappa, S. F., Nespor, M., Ielasi, W., & Miozzo, A. (1997). The representation of stress: Evidence from an aphasic patient. *Cognition*, *65*(1), 1–13. doi:10.1016/S0010-0277(97)00024-3

Caramazza, A. (1997). How many levels of processing are there in lexical access?*Cognitive Neuropsychology*, *14*, 177–208. doi:10.1080/026432997381664

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the 'tip-of-the-tongue' phenomenon. *Cognition*, *64*, 309–343. doi:10.1016/S0010-0277(97)00031-0

Charles-Luce, J. (1985). Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics*, *13*, 309–324.

Cholin, J., Dell, G. S., & Levelt, W. J. M. (2011). Planning and articulation in incremental word production: Syllable-frequency effects in *English*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 109–122. doi:10.1037/a0021322

Cholin, J., & Levelt, W. J. M. (2009). Effects of syllable preparation and syllable frequency in speech production: Further evidence for the retrieval of stored syllables at a post-lexical level. *Language and Cognitive Processes*, *24*, 662–684. doi:10.1080/01690960802348852

Cholin, J., Levelt, W. J. M., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*, 205–235. doi:10.1016/j.cognition.2005.01.009

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*, *98*, 105–135. doi:10.1016/j.cognition.2004.10.006

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, *9*, 287–299.

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*, 214–230. doi:10.1016/S0749-596X(03)00060-3

Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 987–1003. doi:10.1037/0096-1523.18.4.987

Coppock, E. (2010). Parallel grammatical encoding in sentence production: Evidence from syntactic blends. *Language and Cognitive Processes*, *25*(1), 38–49. doi:10.1080/01690960902840261

Costa, A., Caramazza, A., Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1283–1296. doi:10.1037/0278-7393.26.5.1283

Costa, A., Kovacic, D., Fedorenko, E., & Caramazza, A. (2003). The gender congruency effect and the selection of free-standing and bound morphemes: Evidence form Croatian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1270–1282. doi:10.1037/0278-7393.29.6.1270

Cutler, A. (1980). Errors of stress and intonation. In V. A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 67–86). New York, NY: Academic Press.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*, 318–344. doi:10.1037/0278-7393.25.2.318

Damian, M., Vigliocco, G., & Levelt, W. (2001). Effects of semantic context in the naming of pictures and

words. *Cognition*, *81*, B77–B86. doi:10.1016/S0010-0277(01)00135-4

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321. doi:10.1037/0033-295X.93.3.283

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*, 124–142. doi:10.1016/0749-596X(88)90070-8

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, *104*(1), 123–147. doi:10.1037/0033-295X.104.1.123

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review*, *98*, 604–614. doi:10.1037/0033-295X.98.4.604

Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes*, *23*, 583–608. doi:10.1080/01690960801920735

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, *20*, 611–629. doi:10.1016/S0022-5371(81)90202-4

Dell, G. S., Schwartz, M., Martin, M., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838. doi:10.1037/0033-295X.104.4.801

Dhooge, E., & Hartsuiker, R. J. (2010). The distractor frequency effect in picture-word interference: Evidence for response exclusion. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*, 878–891. doi:10.1037/a0019128

Dhooge, E., & Hartsuiker, R. J. (2011). The distractor frequency effect in a delayed picture-word interference task: Further evidence for a late locus of distractor exclusion. *Psychonomic Bulletin & Review*, *18*(1), 116–122. doi:10.3758/s13423-010-0026-0

Dhooge, E., & Hartsuiker, R. J. (2012). Lexical selection and verbal self-monitoring: Effects of lexicality, context, and time pressure in picture-word interference. *Journal of Memory and Language*, *66*, 163–176. doi:10.1016/j.jml.2011.08.004

Fay, D. (1980). Transformational errors. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. New York: Academic Press.

Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, *8*, 505–520.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, *33*, 715–736. doi:10.1006/jmla.1994.1034

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, *35*, 724–755. doi:10.1006/jmla.1996.0038

Ferreira, V. S., & Griffin, Z. M. (2003). Phonological influences on lexical (mis)selection. *Psychological Science*, *14*(1), 86–90. doi:10.1111/1467-9280.01424

Ferreira, V. S., & Humphreys, K. R. (2001). Syntactic influences on lexical and morphological processing in language production. *Journal of Memory and Language*, *44*(1), 52–80. doi:10.1006/jmla.2000.2739

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84. doi:10.1006/jmla.2001.2797

Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, *32*, 669–692. doi:10.1023/A:1026146332132

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). Hillsdale, NJ: Erlbaum.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*, 27–52. doi:10.2307/412187

Fromkin, V. A. (1976). Putting the emPHAsis on the wrong sylLABle. In L. Hyman (Ed.), *Studies in stress and accent* (pp. 15–26). Los Angeles: University of Southern California.

Galante, E., Tralli, A., Zuffi, M., & Avanzi, S. (2000). Primary progressive aphasia: A patient with stress assignment impairment in reading aloud. *Italian Journal of Neurological Science*, *21*(1), 39–48. doi:10.1007/s100720070117

Gann, T. M., & Barr, D. J. (this issue). Speaking from experience: Audience design as expert performance. *Language and Cognitive Processes*.

Garnham, A., Shillcock, R. C., Brown, G. D. A., Mill, A. I. D., & Cutler, A. (1981). Slips of the tongue in the London-Lund corpus of spontaneous conversation. *Linguistics*, *19*, 805–817. doi:10.1515/ling.1981.19.7-8.805

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–177). New York, NY: Academic Press.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 231–255). Amsterdam: North Holland.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production (vol. I): Speech and talk* (pp. 177–220). New York, NY: Academic Press.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.

Garrett, M. F. (1984). The organization of processing structure for language production: Application to aphasic speech. In D. Caplan, A. R. Lecours, & A. Smith (Eds.), *Biological perspectives on language*. Cambridge: MA: MIT Press.

Garrett, M. F. (1992). Disorders of lexical selection. *Cognition*, *42*, 143–180. doi:10.1016/0010-0277(92)90042-G

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*, 613–656. doi:10.1080/016909697386646

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*(1), 1–23.

Gennari, S. P., Mirkovic, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology*, *65*, 141–176. doi:10.1016/j.cogpsych.2012.03.002

Glaser, W. R., & Düngelhoff, F. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 640–654. doi:10.1037/0096-1523.10.5.640

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, *21*, 649–683. doi:10.1080/01690960500181332

Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition*, *103*, 386–412. doi:10.1016/j.cognition.2006.05.010

Gollan, T. H., & Acenas, L.-A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 246–269. doi:10.1037/0278-7393.30.1.246

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, *82*, B1–B14. doi:10.1016/S0010-0277(01)00138-X

Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274–279. doi:10.1111/1467-9280.00255

Griffin, Z. M., & Spieler, D. H. (2006). Observing the what and when of language production for different age groups by monitoring speakers' eye movements. *Brain and Language*, *99*, 272–288. doi:10.1016/j.bandl.2005.08.003

Hantsch, A., Jescheniak, J. D., & Schriefers, H. (2005). Semantic competition between hierarchically related words during speech planning. *Memory and Cognition*, *33*, 984–1000. doi:10.3758/BF03193207

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science*, *8*, 191–219. doi:10.1207/s15516709cog0803_1

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes*, *8*, 291–309. doi:10.1080/01690969308406957

Harley, T. A., & MacAndrew, S. B. G. (1995). Interactive models of lexicalization: Some constraints from speech error, picture naming, and neuropsychological data. In J. P. Levy, D. Bairaktaris, & J. A. Bullinaria (Eds.) *Connectionist models of memory and language* (pp. 311–331). London: UCL Press.

Harley, T. A., & MacAndrew, S. B. G. (2001). Constraints upon word substitution speech errors. *Journal of Psycholinguistic Research*, *30*, 395–418. doi:10.1023/A:1010421724343

Hartley, T. A., & Houghton, G. (1996). A linguistically-constrained model of short-term memory for nonwords. *Journal of Memory and Language*, *35*(1), 1–31. doi:10.1006/jmla.1996.0001

Hartsuiker, R. J., Antón-Mendez, I., & Van Zee, M. (2001). Object attraction in subject agreement construction. *Journal of Memory and Language*, *45*, 546–572. doi:10.1006/jmla.2000.2787

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*, B27–B39. doi:10.1016/S0010-0277(99)00080-3

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, *1*(3), 213–229. doi:10.1017/S1366728998000364

Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's dis- ease: Failure of access or degraded knowledge. *Neuropsychologia*, *30*, 301–314. doi:10.1016/0028-3932(92)90104-T

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*, 464–482. doi:10.1016/j.cognition.2005.02.006

Hwang, H., & Kaiser, E. (in press). Having a syntactic choice is not always better: Syntactic flexibility in Korean production. *Language, Cognition, & Neuroscience*.

Janssen, D. P., Roelofs, A., & Levelt, W. J. (2002). Inflectional frames in language production. *Language and Cognitive Processes*, *17*(3), 209–236.

Janssen, N., & Caramazza, A. (2003). The selection of closed-class words in noun phrase production: The case of Dutch determiners. *Journal of Memory and Language*, *48*, 635–652. doi:10.1016/S0749-596X(02)00531-4

Janssen, N., Schiller, N. O., & Alario, F.-X. (this issue). The selection of closed-class elements during language production: Comments on Jescheniak, Schriefers & Lemhöfer (in press). *Language and Cognitive Processes*.

Jescheniak, J. D., Hahne, A., Hoffmann, S., &Wagner, V. (2006). Phonological activation of category coordinates during speech planning is observable in children but not in adults: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 373–386. doi:10.1037/0278-7393.32.3.373

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1256–1274. doi:10.1037/0278-7393.24.5.1256

Jescheniak, J. D., Schriefers, H., & Lemhoefer, K. (this issue_a). Selection of freestanding and bound gender-marking morphemes in speech production – A review. *Language and Cognitive Processes*.

Jescheniak, J. D., Schriefers, H., & Lemhoefer, K. (this issue_b). What data from the simple-picture naming task tell us about lexical competition – A rejoinder to Janssen, Schiller, and Alario (2012). *Language and Cognitive Processes*.

Kaschak, M. P., Kutta, T. J., & Coyle, J. M. (this issue). Long and short term cumulative structural priming effects. *Language and Cognitive Processes*.

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*(3), 381–388.

Kempen, G., & Harbusch, K. (2004). A Corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann (Ed.), *Language production* (pp. 173–181). Berlin: Mouton.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*, 201–258. doi:10.1207/s15516709cog1102_5

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, *66*, 143–162. doi:10.1016/j.jml.2011.08.003

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*(1), 15–47. doi:10.1016/0010-0277(72)90028-5

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, *128*(3), 416–430.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174. doi:10.1006/jmla.1994.1008

Kuipers, J.-R., & La Heij, W. (2009). The limitations of cascading in the speech production system. *Language and*

Cognitive Processes, *24*(1), 120–135. doi:10.1080/01690960802234177

Kulke, F., & Blanken, G. (2001). Phonological and syntactic influences on semantic misnamings in aphasia. *Aphasiology*, *15*(1), 3–15. doi:10.1080/02687040042000070

Laganaro, M. (2005). Syllable frequency effect in speech production: Evidence from aphasia. *Journal of Neurolinguistics*, *18*, 221–235. doi:10.1016/j.jneuroling.2004.12.001

Laganaro, M. (2008). Is there a syllable frequency effect in aphasia or in apraxia of speech or both? *Aphasiology*, *22*, 1191–1200. doi:10.1080/02687030701820469

Laganaro, M., & Alario, F.-X. (2006). On the locus of the syllable frequency effect in language production. *Journal of Memory and Language*, *55*, 178–196. doi:10.1016/j.jml.2006.05.001

Laganaro, M., Vacheresse, F., & Frauenfelder, U. H. (2002). Selective impairment of lexical stress assignment in an Italian-speaking aphasic patient. *Brain and Language*, *81*, 601–609. doi:10.1006/brln.2001.2550

Laine, M., & Martin, N. (1996). Lexical retrieval deficit in picture naming: Implications for word production models. *Brain and Language*, *53*, 283–314. doi:10.1006/brln.1996.0050

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*, 1–22. doi:10.1016/0010-0277(92)90038-J

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 13464–13471. doi:10.1073/pnas.231459498

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, *98*(1), 122–142. doi:10.1037/0033-295X.98.1.122

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*, 239–269. doi:10.1016/0010-0277(94)90030-2

Lindsley, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*, *7*(1), 1–19. doi:10.1016/0010-0285(75)90002-X

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory and Cognition*, *7*, 485–495. doi:10.3758/BF03198265

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703. doi:10.1037/0033-295X.101.4.676

Mackay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, *89*, 483–506. doi:10.1037/0033-295X.89.5.483

McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, *25*, 188–230. doi:10.1006/cogp.1993.1005

McMillan, C. T., & Corley, M. (2010). Cascading influences on the production of speech: Evidence from articulation. *Cognition*, *117*, 243–260. doi:10.1016/j.cognition.2010.08.019

McMillan, C. T., Corley, M., & Lickley, R. J. (2009). Articulatory evidence for feedback and competition in speech production. *Language and Cognitive Processes*, *24*(1), 44–66. doi:10.1080/01690960801998236

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312. doi:10.1006/jmla.1997.2543

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*, 503–535. doi:10.1037/0278-7393.33.3.503

Marx, E. (1999). Gender processing in speech production: Evidence from German speech errors. *Journal of Psycholinguistic Research*, *28*, 601–621. doi:10.1023/A:1023216927241

Melinger, A., & Abdel Rahman, R. (2013). Lexical selection is competitive: Evidence from indirectly activated semantic associates during picture naming. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*, 348–364. doi:10.1037/a0028941

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition*, *42*, 181–211. doi:10.1016/0010-0277(92)90043-H

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of Memory and Language*, *35*, 477–496. doi:10.1006/jmla.1996.0026

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131–147. doi:10.1016/S0749-596X(02)00509-0

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*, B25–B33. doi:10.1016/S0010-0277(98)00009-2

Miceli, G., & Caramazza, A. (1993). The assignment of word stress in oral reading: Evidence from a case of acquired dyslexia. *Cognitive Neuropsychology*, *10*, 273–295. doi:10.1080/02643299308253465

Miozzo, M., & Caramazza, A. (1997). Retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1410–1423. doi:10.1037/0278-7393.23.6.1410

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 555–563. doi:10.1037/0278-7393.28.3.555

Motley, M. T., & Baars, B. J. (1976). Semantic bias effects on the outcomes of verbal slips. *Cognition*, *4*, 177–187. doi:10.1016/0010-0277(76)90003-2

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: The case of English and Russian. *The Quarterly Journal of Experimental Psychology*, *66*(8), 1601–1619.

Navarette, E., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. *Journal of Memory and Language*, *53*, 359–377. doi:10.1016/j.jml.2005.05.001

Navarrete, E., Del Prato, P., & Mahon, B. Z. (2012). Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology*, *3*, 38.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative

semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252.
Oppermann, F., Jescheniak, J. D., Schriefers, H., & Görges, F. (2010). Semantic relatedness among objects promotes the activation of multiple phonological codes during object naming. *The Quarterly Journal of Experimental Psychology*, *63*, 356–370. doi:10.1080/17470210902952480
Perret, C., Schneider, L., Dayer, G., & Laganaro, M. (this issue). Convergences and divergences between neurolinguistic and psycholinguistic data in the study of phonological and phonetic encoding: A parallel investigation of syllable frequency effects in brain-damaged and healthy speakers. *Language and Cognitive Processes*.
Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 539–557. doi:10.1037/0278-7393.24.3.539
Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651. doi:10.1006/jmla.1998.2592
Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, *134*(3), 427.
Port, R., & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, *13*, 455–471.
Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, *42*, 168–182. doi:10.1006/jmla.1999.2668
Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*, 460–499. doi:10.1037/0033-295X.107.3.460
Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142. doi:10.1016/0010-0277(92)90041-F
Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284. doi:10.1016/S0010-0277(97)00027-9
Rosinski, R. R. (1977). Picture-word interference is semantically based. *Child Development*, *48*, 643–647. doi:10.2307/1128667
Schiller, N. O. (1998). The effect of visually masked syllable primes on the naming latencies of words and pictures. *Journal of Memory and Language*, *39*, 484–507. doi:10.1006/jmla.1998.2577
Schiller, N. O., Fikkert, P., & Levelt, C. C. (2004). Stress priming in picture naming: An SOA study. *Brain and Language*, *90*, 231–240. doi:10.1016/S0093-934X(03)00436-X
Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199–227.
Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time-course of lexical access in production: Picture-word interference studies. *Journal of Memory and Language*, *29*, 86–102. doi:10.1016/0749-596X(90)90011-N
Sevald, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition*, *53*(2), 91–127. doi:10.1016/0010-0277(94)90067-1
Shallice, T., & McGill, J. (1978). The origin of mixed errors. In J. Requin (Ed.), *Attention and performance VII* (pp. 193–208). Hillsdale, NJ: Erlbaum.
Slowiaczek, L., & Dinnsen, D. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, *13*, 325–341.
Smith, M., & Wheeldon, L. (1999). High-level processing scope in spoken sentence production. *Cognition*, *73*, 205–246. doi:10.1016/S0010-0277(99)00053-0
Smith, M. C., & Wheeldon, L. R. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*, 675–686. doi:10.1037/0278-7393.30.3.675
Stallings, L. M., MacDonald, M. C., & O'Seaghdha, P. G. (1998). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in Heavy-NP shift. *Journal of Memory and Language*, *39*, 392–417. doi:10.1006/jmla.1998.2586
Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 686–698. doi:10.1037/0278-7393.21.3.686
Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 869–918.
Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, *60*, 308–327. doi:10.1016/j.jml.2008.11.002
Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, *114*, 447–454. doi:10.1016/j.cognition.2009.11.003
Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (pp. 143–186). London: Erlbaum.
Tanaka, M. N., Branigan, H. P., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language*, *65*, 318–330. doi:10.1016/j.jml.2011.04.009
Turner, E., & Rommetveit, R. (1967). Experimental manipulation of the production of active and passive voice in children. *Language and Speech*, *10*, 169–180.
Underwood, G. (1976). Semantic interference from unattended printed words. *British Journal of Psychology*, *67*, 327–338. doi:10.1111/j.2044-8295.1976.tb01519.x
Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*, 225–258.
Van Nice, K. Y., & Dietrich, R. (2003). Task sensitivity of animacy effects: Evidence from German picture descriptions. *Linguistics*, *41*, 825–849.
Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, *8*, 314–317. doi:10.1111/j.1467-9280.1997.tb00444.x
Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, *34*, 186–215. doi:10.1006/jmla.1995.1009
Vigliocco, G., Vinson, D. P., Indefrey, P., Levelt, W. J. M., & Hellwig, F. (2004). Role of grammatical gender and semantics in German word production. *Journal of Experimental Psychology: Learning Memory, and Cognition*, *30*, 483–497. doi:10.1037/0278-7393.30.2.483
Vigliocco, G., Vinson, D. P., & Paganelli, F. (2004). Grammatical gender and meaning. *In Proceedings of the 26th annual meeting of the cognitive science society* (pp. 1405–1410), Chicago.
Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Oscillator-based control of the serial ordering of phonology

in speech production. *Cognitive Psychology*, *41*, 101–175. doi:10.1006/cogp.2000.0739

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 423–440. doi:10.1037/a0018619

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, *32*, 251–276. doi:10.1016/S0095-4470(03)00032-9

Wasow, T. (1997a). Remarks on grammatical weight. *Language Variation and Change*, *9*(1), 81–105. doi:10.1017/S0954394500001800

Wasow, T. (1997b). End-weight from the speaker's perspective. *Journal of Psycholinguistic Research*, *26*, 347–361. doi:10.1023/A:1025080709112

Yamashita, H., & Chang, F. (2001). "Long before short" preference in the production of a head-final language. *Cognition*, *81*, B45–B55. doi:10.1016/S0010-0277(01)00121-4